View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Elsevier - Publisher Connector

Chaotic Relaxation*

D. CHAZAN AND W. MIRANKER
IBM Watson Research Center, Yorktown Heights, New York

Communicated by Alan J. Hoffman

ABSTRACT

In this paper we consider relaxation methods for solving linear systems of equations. These methods are suited for execution on a parallel system of processors. They have the feature of allowing a minimal amount of communication of computational status between the computers, so that the relaxation process, while taking on a chaotic appearance, reduces programming and processor time of a bookkeeping nature. We give a precise characterization of chaotic relaxation, some examples of divergence, and conditions guaranteeing convergence.

INTRODUCTION

In this paper we discuss a preliminary study of relaxation methods for solving linear systems of equations. These methods are of a form capable of execution on a set of parallel processors. In the most general case, there is a minimum of communication of computational status between the computers, so that the relaxation process takes on a chaotic appearance.

Let the linear system be given by

$$Ax = d, \tag{0.1}$$

where x is an n vector. If A is symmetric and positive definite, solving (1) by relaxation may be described as successive univariate minimization, usually in coordinate directions, of the quadratic form

$$Q(x) = \tfrac{1}{2}(x, Ax) - (d, x). \tag{0.2}$$

* Dedicated to Professor A. M. Ostrowski on his 75th birthday.

Copyright © 1969 by American Elsevier Publishing Company, Inc.

For the Gauss-Seidel method a univariate minimization is completed prior to commencing any other. Now we suppose that r processors are available for this computation. Each may be sequentially assigned to effect a minimization. However, during the time that a processor is determining what displacement at a point x in n-space is to be effected, other processors may complete their tasks and displace the point x. Thus, the displacement to a minimum determined by a processor at the point x may be made at some other point. This apparently arbitrary pattern is a familiar one. For the Jacobi scheme n univariate minimization displacements are computed at a point x, but only one of the n displacements occurs at x.

In the general chaotic case, the r processors will rapidly get out of phase, since they themselves may be of different quality or the specific relaxations performed may be of varying difficulty.

The problem of chaotic relaxation was suggested to the authors by J. Rosenfeld who has conducted extensive numerical experiments with chaotic relaxation [3]. Rosenfeld found that the use of more processors decreased the computation time for solving the linear system. The precise form of the decrease depended on the relative number of processors, r, to n, the dimension. The chaotic form of the relaxation eliminated considerable programming and computer time in coordinating the processors and the algorithm. His experiments also exhibited a diminution in the amount of overrelaxation allowable for convergence in a chaotic mode.

Chaotic relaxation is related to free steering which has been studied by A. Ostrowski [6] and S. Schechter [5]. It is however more general because not only is the order of relaxation free but the antecedent values upon which a relaxation is based are chosen freely component by component from a number of predecessor states. In [6], Ostrowski obtained his free steering results under the condition that the matrix A be a so-called H matrix. Schechter [5] considered free block relaxation and obtained convergence for so-called monotone A. For simplicity, we have not considered block relaxation, although the methods we use may be extended to obtain convergence statements for chaotic block relaxation. When the chaotic relaxation is specialized to ordinary free steering, our results recover some of the statements of Ostrowski.

In Sect. 1 chaotic relaxation is defined along with a consistency notion. In addition, we give several examples and introduce an important subclass of chaotic schemes called periodic schemes. In Section 2 we give a variational characterization to chaotic schemes and two examples in the periodic case which diverge. In Sect. 3 periodic schemes are reduced to a difference equation in an augmented state space. The difference equation is then used to obtain two convergence theorems. In Sect. 4 we give another convergence result in the periodic case by exploiting the variational characterization derived in Sect. 2. Finally in Sect. 5, we turn to the fully chaotic case and give our main theorem. This theorem gives necessary and sufficient conditions for a chaotic iteration scheme to converge. A number of corollaries are then given which produce classes of matrices for which chaotic relaxation converges.

1. DEFINITION OF TERMS

In this section, we define a chaotic iteration scheme. We then introduce the notion of consistency of a chaotic scheme with a linear system. We then give several examples of chaotic schemes. One of these examples is a class of schemes, called periodic schemes, which we will study in subsequent sections.

DEFINITION. A chaotic iterative scheme is a class of sequences of n-vectors, x^j, $j = 1, \ldots$. Each sequence in this class is defined recursively by

$$x_i^{j+1} = \begin{cases} x_i^j, & i \neq k_{n+1}(j) \\ \displaystyle\sum_{\alpha=1}^{n} b_\alpha^i x_\alpha^{j-k_\alpha(j)} + (c^i, d), & i = k_{n+1}(j). \end{cases} \tag{1.1}$$

The initial n-vector x^1 and the n-vector d are arbitrary. The mapping (2.1) depends on a triple $(B, C, \mathscr{S})$. Here B and C are $n \times n$ matrices. The ith rows of B and C are denoted respectively by $b^i = (b_i^1, \ldots, b_i^n)$ and by c^i. $\mathscr{S}$ is a sequence of $n+1$-vectors; $\mathscr{S} = \{k_1(j), \ldots, k_{n+1}(j)\}$, $j = 1, \ldots,$ with the following properties:

For some fixed integer $s > 0$

(i) $\quad 0 \leqslant k_i(j) < s, \qquad i = 1, \ldots, n, \qquad j = 1, \ldots .$

(ii) $\quad 1 \leqslant k_{n+1}(j) \leqslant n, \qquad j = 1, \ldots .$

Moreover, $k_{n+1}(j) = i$ infinitely often for each i, $s \leqslant i \leqslant n$.

We will identify a chaotic relaxation scheme by the triple $(B, C, \mathscr{S})$, with the implication that once B, C, and $\mathscr{S}$ are fixed a particular sequence of the class is chosen.

This definition may be interpreted as follows: At each instant of time j, the $k_{n+1}(j)$th component of x^j is updated while the remaining $n-1$ components are unchanged. The updating uses the first component of $x^{j-k_1(j)}$, the second component of $x^{j-k_2(j)}$, etc. Every component is updated infinitely often, and no update uses a value of a component which was produced by an update s or more steps previously.

Since we will use chaotic iteration procedures to solve the linear system $Ax = d$, we introduce a consistency notion in the following

DEFINITION. A chaotic iteration procedure is consistent with a linear system $Ax = d$, with solution z, if z is a fixed point of the mapping (1.1), i.e.,

$$z = Bz + Cd. \tag{1.2}$$

Notice that (1.2) is independent of $\mathscr{S}$.

From now on we will assume that the chaotic scheme is consistent with the linear system $Ax = d$. We note then that if $y^j = x^j - z$, then the y^j are given by the following recursive scheme

$$y_i^{j+1} = \begin{cases} y_i^j, & i \neq k_{n+1}(j) \\ \sum\limits_{\alpha} b_\alpha^i y_\alpha^{j-k_\alpha(j)}, & i = k_{n+1}(j). \end{cases} \tag{1.3}$$

(1.3) is a chaotic iteration scheme corresponding to the homogeneous system $Ax = 0$. Since we are only considering consistent schemes, we may set $d = 0$ without loss of generality.

Obtaining Consistent Schemes.

A consistent chaotic scheme may be obtained by choosing B and C as follows. Let D be the diagonal matrix whose iith element is a_{ii} and let $E = D - A$. Then the choosing $B = D^{-1}E$ and $C = D^{-1}$ produces a consistent scheme. A larger class of consistent schemes correspond to over and under relaxation with parameter ω. We denote these schemes by $(\mathscr{S}, B^\omega, C^\omega)$. Here $B^\omega = I - \omega D^{-1}A$ and $C^\omega = \omega D$. Note that $B^1 = B$ and $C^1 = C$.

We will now give several examples of chaotic schemes.

Example 1 (Gauss-Seidel). Let $\mathscr{S}$ be defined by $k_1(j) = k_2(j) = \cdots = k_n(j) = 0$ and $k_{n+1}(j) \equiv (j-1)(\text{mod } n)$. Here $s = 1$ and $k_{n+1}(j) = i$, for each i, $i = 1, \ldots, n$ exactly once in every n updates. This scheme is precisely the Gauss-Seidel relaxation procedure.

Example 2 (Jacobi). Let $\mathscr{S}$ be defined by $k_1(j) = \cdots = k_n(j) = k_{n+1}(j)$ and $k_{n+1}(j) \equiv (j-1)(\operatorname{mod} n)$. Here $s = n$ and $k_{n+1}(j) = i$, for each i, $i = 1, \ldots, n$ exactly once in every n updates. This scheme is precisely the Jacobi relaxation procedure.

Example 3 (No Lapping Schemes). We introduce the no lapping convention. Here a number of processors are at work. Each processor as it becomes available is assigned sequentially to relax components. At time j a processor P is assigned to relax the component x_i say. We imagine that it takes so long to perform this calculation that each of the components other than x_i are either updated or assigned to be updated while P is still calculating to relax the component x_i. Ultimately, a newly available processor would be assigned to update x_i. The no lapping convention rules this out and stops all assignments until the component at x_i has been updated by processor P. For this procedure we see that $s = 2n - 1$ and again all components are relaxed infinitely often.

Periodic Schemes

We now introduce an important subclass of schemes which we call periodic schemes. These schemes are suitable for parallel computation but do not correspond to any of the classical methods. This class of schemes corresponds to the case where $k_{n+1}(j)$ is a periodic function of j. Periodic schemes are an orderly version of the chaotic case. The processors compute simultaneously but in such a regular fashion that a repeating periodic pattern exists in the computation. Formally, a periodic scheme is given by the following definition.

DEFINITION. Corresponding to each integer* r, $1 \leqslant r \leqslant n$, a periodic scheme is a chaotic scheme with $k_1(j) = \cdots = k_n(j)$, and

$$k_1(j) = \begin{cases} r-1, & j \geqslant r-1 \\ j-1, & j < r-1 \end{cases},$$

and with

$$k_{n+1}(j) \equiv (j-1)(\operatorname{mod} n).$$

A periodic scheme may be described by a dependency graph whose vertices are the integers, $j = 1, 2, \ldots$.

* r is the number of processors assigned to the relaxation calculation.

The axis is indexed by j. At the instant j, the vector x^j is viewed as being written into memory. Each curve in the graph traces the activity of a particular processor. For $j_1 < j_2$, j_2 can be reached from j_1 along a trajectory in the graph, if and only if the updated component of x^{j_2} (i.e., $x^{j_2}_{k_{n+1}(j_2-1)}$) depends on the j_1st iterate formally (in the sense that the right-hand side of 1.1 contains a term with superscript j_1). (Alternatively, in terms of the parallel processors, this happens if and only if x^{j_1} and x^{j_2} are successive updates of the same processor.) The graph for $r = 3$ and $n = 6$ is shown in Fig. 1.1 where the label from j_1 to j_2 is $k_{n+1}(j_2 - 1)$.

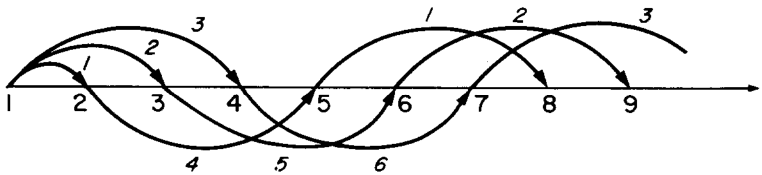

FIG. 1.1. Graph of periodic case with three processors and dimension 6.

Each time a vector is written, j is increased by unity, although in reality in the computing system, the actual interval of time between which updated components are written into memory will vary from essentially zero to some positive number in a more or less unpredictable manner.

Conventions

Unless otherwise specified, the remainder of this paper deals with A symmetric and $A = MM^*$. The columns of M are m^i, $i = 1, \ldots, n$. In addition, D is the diagonal matrix $(a_{ij}\delta_{ij})$, $E = D - A$, $B = D^{-1}E$, $C = D^{-1}$, $B^{\omega} = I - \omega D^{-1}A$, and $C^{\omega} = \omega D^{-1}$. In Sect. 5, A will not be required to be symmetric.

2. TWO DIVERGENT EXAMPLES

In this section we will derive a representation of chaotic relaxation as a sequence of minimizations of a quadratic form. Then we will use this representation to explore two divergent examples which help to characterize the difficulties in studying chaotic relaxation.

A variational characterization of chaotic relaxation in the periodic case is given in the following theorem.

THEOREM 2.1. *Let x^j be a sequence of iterates produced by a periodic scheme parametrized by r and ω. Let $y^j = Mx^j$. If m^i is the ith column of M, then*

$$y^{j+1} = y^j - \omega\alpha^j m^{k_{n+1}(j)} \tag{2.1}$$

where α^j minimizes

$$\|y^{j-r+1} - \alpha^j m^{k_{n+1}(j)}\|. \tag{2.2}$$

Here $\|\ \|$ denotes the Euclidean norm.

Proof. We will first set $\omega = 1$. In this case we have

$$x_i^{j+1} = \begin{cases} x_i^j, & i \neq k_{n+1}(j) \\ (b^{k_{n+1}(j)}, x^{j-r+1}), & i = k_{n+1}(j). \end{cases} \tag{2.3}$$

Let v^j be the vector whose ith component is δ_{ij}. Then (2.3) may be written as

$$\begin{aligned} x^{j+1} &= x^j - x^j_{k_{n+1}(j)} v^{k_{n+1}(j)} + (v^{k_{n+1}(j)}, Bx^{j-r+1}) v^{k_{n+1}(j)} \\ &= x^j - \alpha^j v^{k_{n+1}(j)}. \end{aligned} \tag{2.4}$$

Now

$$x^j_{k_{n+1}(j)} = x^{j-r+1}_{k_{n+1}(j)} \tag{2.5}$$

since $k_{n+1}(i) \neq k_{n+1}(j)$, $j - n < i < j$ (i.e., no component is updated twice before each of the $n - 1$ remaining ones are updated once). Using (2.5) in (2.4) we may write α^j as

$$\alpha^j = (v^{k_{n+1}(j)}, (I - B)x^{j-r+1}). \tag{2.6}$$

When $\omega \neq 1$ it is easy to verify that α^j is replaced by $\omega\alpha^j$. It is also easy to verify that $\alpha^j v^{k_{n+1}(j)}$ is the displacement along $v^{k_{n+1}(j)}$ which minimizes

$$(x^{j-r+1} - \alpha^j v^{k_{n+1}(j)}, A(x^{j-r+1} - \alpha^j v^{k_{n+1}(j)})). \tag{2.7}$$

In terms of y^j (2.4) and (2.7) become respectively

$$y^{j+1} = y^j - \alpha^j m^{k_{n+1}(j)} \tag{2.8}$$

$$\|y^{j-r+1} - \alpha^j m^{k_{n+1}(j)}\|^2. \tag{2.9}$$

This concludes the proof of the theorem.

We note here that α^j can thus be completely determined from knowledge of y^{j-r+1}.

We now consider two divergent examples.

Example 1. Let

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$M = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad m^i = (1, 0, 0)^T, \qquad i = 1, 2, 3.$$

Let

$$r = 2, \qquad \omega = 1 + \varepsilon.$$

Then

$$y^{j+1} = y^j - \omega\alpha^j m^{k_{n+1}(j)}.$$

Since $y^j = Mx^j$, y^j has the form $y^j = (z^j, 0, 0)$. Thus, the iteration scheme is described equivalently by the following iteration scheme for the scalar z^j,

$$z^{j+1} = z^j - \omega\alpha^j.$$

Using (2.2), this can be written as

$$z^{j+1} = z^j - \omega z^{j-1}. \tag{2.10}$$

The eigenvalues of this second-order iterative scheme may be easily shown to be no smaller than 1 if $\varepsilon > 0$. It is instructive to consider Fig. 2.1 where the evolution of a solution of the difference equation (2.10) is graphically illustrated. The solution has initial values $z^1 = z^2 = -1$ and blows up exponentially. z^j is located at the point labeled j.

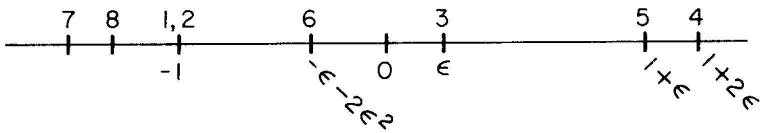

FIG. 2.1. Evolution of solution of Eq. (3.1). Iteration number is labeled above the axis and iterate values below.

Example 2. Let A and M be as above. Let $\omega = 1$ and $r = 3$. Again writing everything in terms of z^i we obtain

$$z^{i+1} = z^i - z^{i-2}. \tag{2.11}$$

This iterative scheme also has eigenvalues greater than one. The way the process (2.11) blows up is illustrated in Fig. 2.2.

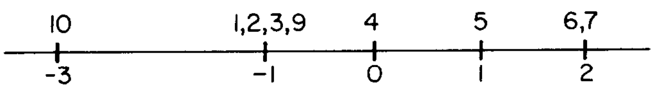

FIG. 2.2. Evolution of solution of Eq. (3.3). Iteration number is labeled above the axis and iterate values below.

These examples show that underrelaxation ($\omega < 1$) must be used with the matrix A given here. The examples also point out the reasons why both of these processes are unstable. Since we view each iteration or relaxation step as a motion in the direction m^i, we can say that in the examples given here the directions of motion were so close together that tasks were duplicated. It seems as though the event when $m^1 = m^2 = \cdots = m^n$ is, in a sense, the worst case for such schemes. A conjecture is therefore that if ω is such that the schemes converge for the specific A above, they converge for all A. In particular, $\omega < 1$ should be sufficient to guarantee convergence for the two-processor ($r = 2$) case. None of these questions has been settled.

The apparently worst case ($m^1 = m^2 = \cdots = m^n$) corresponds to a singular matrix. It seems clear that there exist nonsingular matrices which are arbitrarily close to the worst case form and which exhibit divergence.

3. A DIFFERENCE EQUATION AND CONVERGENCE IN THE PERIODIC CASE

In this section we will derive a representation for the periodic scheme, as a first-order difference equation in Euclidean $n + r - 1$-space. We illustrate this for $r = 2$. The general case is quite similar. We then present two cases for which convergence is demonstrated.

Note that for the periodic scheme $k_{n+1}(j) = (j - 1) \bmod n$ so that $k_{n+1}(j - 1) \neq k_{n+1}(j)$ for $i = 1, \ldots, n - 1$. Thus, $x^j_{k_{n+1}(j)} = x^{j-1}_{k_{n+1}(j)} = \cdots = x^{j-(n-1)}_{k_{n+1}(j)}$. Schematically, this fact can be represented by the graph below (for $n = 7$): where the vertices of the graph are labeled (i, j).

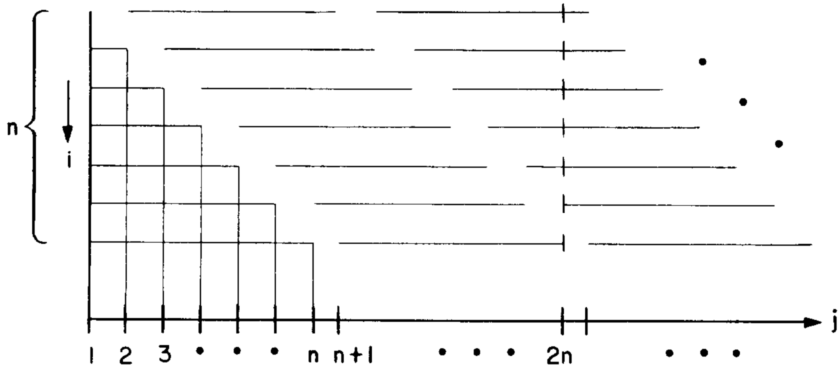

FIG. 3.1. Graph of the periodic case for $r = 2$ and $n = 7$.

(i_1, j_1) is connected to (i_2, j_2) if $i_1 = i_2$ and no update of x_{i_1} occurs for j with $j_1 < j \leqslant j_2$. Referring to this graph it is easy to see that for $r = 2$, the updating is given by the following set of relations (3.1).

$$
\begin{aligned}
x_1^{(q+1)n+2} &= \langle b^1, x^{(q+1)n} \rangle = \sum_{\gamma=1}^{n-1} b_\gamma{}^1 x_\gamma{}^{qn+1+\gamma} + b_n{}^1 x_n{}^{(q+1)n} \\
x_2^{(q+1)n+3} &= \langle b^2, x^{(q+1)n+1} \rangle = \sum_{\gamma=1}^{n} b_\gamma{}^2 x_\gamma{}^{qn+1+\gamma} \\
x_3^{(q+1)n+4} &= \langle b^3, x^{(q+1)n+2} \rangle = b_1{}^3 x_1{}^{(q+1)n+2} + \sum_{\gamma=2}^{n} b_\gamma{}^3 x_\gamma{}^{qn+1+r} \\
&\vdots \\
x_n^{(q+1)n+1+n} &= \langle b^n, x^{(q+1)n+n-1} \rangle = \sum_{\gamma=1}^{n-2} b_\gamma{}^n x_\gamma{}^{(q+1)n+1+\gamma} \\
&\quad + b^n_{n-1} x^{(q+1)n+n-1}_{n-1} + b_n{}^n x_n{}^{(q+1)n+n}
\end{aligned} \tag{3.1}
$$

where b^i is the ith row of B and q is any positive integer.

We will explain how the first equation in (3.1) comes about. A horizontal line segment in the Fig. 3.1 represents a computed value of a component. The segment begins when that computed value is written into memory. The segment terminates one unit of index prior to the next write into memory of that component. When a processor is assigned to effect a relaxation, the memory contents it has available is determinable by drawing a vertical line at the value j of the assignment. Those horizontal line segments which this vertical line meets are representative of the component value currently in memory; hence in the present case, $r = 2$ processors, the first component receives for its computation that processor which itself has just written the penultimate component. The memory contents are then determinable by drawing a vertical line through the left endpoint of a penultimate horizontal line segment. In Fig. 3.1 this vertical line is drawn and the memory contents are thus seen to belong to one group of segments forming a parallelogram of horizontal line segments with the exception of the last component which belongs to a predecessor group. The superscripts in the right member of the first equation in (3.1) fall into these two groups. In a similar manner the remaining equations in (3.1) may be deduced from the figure.

Finally we add the equation

$$x_n^{(q+2)n} = x_n^{(q+1)n+n} = x_n^{qn+1+n}.$$

Let z^q be the n-dimensional vector whose components are

$$z_i^q = x_i^{(q-1)n+1+i}, \qquad i = 1, \ldots, n,$$

and let $w^q = x_n^{qn}$. Multiplying the ith equation in (3.1) by a_{ii}, $i = 1, \ldots, n$, we obtain a recurrence relation for (z^q, w^q):

$$\begin{aligned} Dz^{q+1} &= L_0 z^{q+1} + L_1 z^{q+1} + (L_2 + L_2^*) z^q + L_1^* z^q - a_{1,n} w^q v^1 \\ w^{q+1} &= z_n^q. \end{aligned} \tag{3.2}$$

Here L_0, L_1, L_2 are $n \times n$ matrices whose elements $(L_p)_{i,j}$ are:

$$\begin{aligned} (L_0)_{i,j} &= \begin{cases} -a_{n,1}, & i = n, \quad j = 1 \\ 0, & \text{otherwise} \end{cases} \\ (L_1)_{i,j} &= \begin{cases} 0, & i = n, \quad j = 1 \\ 0, & j \geqslant i - 1 \\ -a_{i,j}, & \text{otherwise} \end{cases} \end{aligned} \tag{3.3}$$

and

$$(L_2)_{i,j} = \begin{cases} a_{i,j}, & i = j - 1 \\ 0, & \text{otherwise.} \end{cases}$$

Our first case for which convergence is obtained is given in the following theorem.

THEOREM 3.1. *If $n = 3$, $\omega = 1$, $r = 2$, and A is symmetric positive definite, then the periodic scheme converges.*

Proof. Without loss of generality we may select m_1, m_2, and m_3 such that $m_1 = (1, 0, 0)$, $m_2 = (w_1, w_2, 0)$, $m_3 = (v_1, v_2, v_3)$ and

$$M = \begin{pmatrix} 1 & w_1 & v_1 \\ 0 & w_2 & v_2 \\ 0 & 0 & v_3 \end{pmatrix},$$

and furthermore we may assume that $w_1^2 + w_2^2 = v_1^2 + v_2^2 + v_3^2 = 1$.

Then $a_{i,j} = \langle m_i, m_j \rangle$ and $a_{i,i} = \langle m_i, m_i \rangle = 1$. If we let $S^q = (z^q, w^q)$, we obtain from (3.2) and (3.3):

$$S^{q+1} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -v_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} S^{q+1} - \begin{pmatrix} 0 & w_1 & 0 & v_1 \\ w_1 & 0 & \langle w, v \rangle & 0 \\ 0 & \langle w, v \rangle & 0 & 0 \\ 0 & 0 & -1 & 0 \end{pmatrix} S^q. \tag{3.4}$$

Reordering the equation through $u_1^q = S_1^q$, $u_2^q = S_3^q$, $u_3^q = S_2^q$, and $u_4^q = S_4^q$, we have:

$$u^{q+1} = -\begin{pmatrix} 0 & 0 & 0 & 0 \\ v_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} u^{q+1} - \begin{pmatrix} 0 & 0 & w_1 & v_1 \\ 0 & 0 & \langle w, v \rangle & 0 \\ w_1 & \langle w, v \rangle & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} u^q.$$

Let

$$P = \begin{pmatrix} 0 & 0 \\ v_1 & 0 \end{pmatrix}, \quad Q = \begin{pmatrix} w_1 & v_1 \\ \langle w, v \rangle & 0 \end{pmatrix}, \quad \text{and} \quad R = \begin{pmatrix} w_1 & \langle w, v \rangle \\ 0 & -1 \end{pmatrix}.$$

Then

$$u^{q+1} + \begin{pmatrix} P & 0 \\ 0 & 0 \end{pmatrix} u^{q+1} + \begin{pmatrix} 0 & Q \\ R & 0 \end{pmatrix} u^q = 0. \tag{3.5}$$

If u is an eigenvector of (3.5), we have

$$\lambda u_1 + \lambda P u_1 + Q u_2 = 0$$

$$\lambda u_2 + R u_1 = 0$$

where u is split into two two-dimensional components u_1 and u_2. Thus,

$$\lambda^2 u_1 + \lambda^2 P u_1 - QR u_1 = 0.$$

Convergence of the scheme will follow if $|\lambda| < 1$. In order to show that $|\lambda| < 1$, it is enough to show that the roots β of

$$|\beta(I + A) - BC| = 0,$$

where $|\ |$ denotes a determinant, are smaller than 1 in magnitude.

Now

$$\beta = \frac{b \pm \sqrt{b^2 - 4c}}{2}$$

where

$$b = (w_1{}^2 + \langle w, v\rangle^2 + v_1{}^2 - v_1 w_1 \langle w, v\rangle) \qquad \text{and} \qquad c = w_1 v_1 \langle w, v\rangle.$$

If $b^2 < 4c$, then $|\beta|^2 = c < 1$. Suppose $b^2 \geqslant 4c$. We must show then that

$$|b \pm \sqrt{b^2 - 4c}| < 2.$$

But

$$b = w_1{}^2 + v_1{}^2 + \langle w, v\rangle w_2 v_2$$

$$= v_1{}^2 + w_1{}^2 + w_1 v_1 w_2 v_2 + (w_2 v_2)^2.$$

Then b is positive since

$$w_1{}^2 + v_1{}^2 \geqslant 2|w_1 v_1| > |(w_1 v_1)(w_2 v_2)|.$$

Also $b \leqslant 2$, since

$$\begin{aligned} b &= w_1{}^2 + v_1{}^2 + (w_2 v_2)\langle w, v\rangle \\ &\leqslant w_1{}^2 + v_1{}^2 + |w_2 v_2| \\ &\leqslant 2 - w_2{}^2 - v_2{}^2 + 2|w_2 v_2| \\ &\leqslant 2 - (|w_2| - |v_2|)^2 \\ &\leqslant 2. \end{aligned}$$

Then to conclude the proof there remains only to show that $(2 - b) > \sqrt{b^2 - 4c}$. Using the positivity of $(2 - b)$ we wish to have

$$4 - 4b + b^2 > b^2 - 4c$$

or

$$1 \geqslant b - c.$$

But

$$\begin{aligned} b - c &= w_1{}^2 + v_1{}^2 + (w_2 v_2)^2 + (w_1 v_1)(w_2 v_2) \\ &\quad - ((w_1 v_1)^2 + (w_1 v_1)(w_2 v_2)) \\ &= w_1{}^2 + v_1{}^2 + (w_2 v_2)^2 - (w_1 v_1)^2 \\ &\leqslant w_1{}^2 + v_1{}^2 + (1 - w_1{}^2)(1 - v_1{}^2) - w_1{}^2 v_1{}^2 \\ &= 1, \end{aligned}$$

which concludes the proof.

Our second case in which we obtain convergence is given in Theorem 3.2.

THEOREM 3.2. *Let* $r = 2$ *and suppose* $a_{1,n} = 0$ *and* $D + L_2 + L_2{}^*$ *and* A *are positive definite. Then the periodic scheme described by* (3.2) *and* (3.3) *converges.*

Proof. The proof was suggested by W. Kahn [1]. A similar proof along lines proposed by P. Stein [2] is also possible.

Since $a_{1,n} = 0$, L_0 vanishes, and the iteration scheme reduces to

$$Dz^{q+1} = L_1 z^{q+1} + (L_2 + L_2^*)z^q + L_1^* z^q.$$

Let λ be the eigenvalue of largest modulus of this scheme and let z be the corresponding eigenvector; i.e.,

$$\lambda Dz = \lambda L_1 z + (L_2 + L_2^*)z + L_1^* z.$$

Then

$$\lambda = \frac{\langle z, (L_2 + L_2^*)z\rangle + \langle z, L_1 z\rangle^*}{\langle z, Dz\rangle - \langle z, L_1 z\rangle}.$$

We abbreviate this expression as follows:

$$\lambda = \frac{\alpha + \beta^*}{\gamma - \beta}.$$

Now $\gamma + \alpha > 0$, since $D + L_2 + L_2^*$ is positive definite. $\gamma - \alpha - \beta - \beta^* > 0$, since A is positive definite. Let $\beta = \beta_1 + i\beta_2$, $\delta = \alpha + \beta_1$, and $\eta = \gamma - \beta_1$. Then

$$|\lambda| = \frac{|\delta - i\beta_2|}{|\eta - i\beta_2|}. \tag{3.6}$$

Since $\eta - \delta > 0$ and $\eta + \delta > 0$, then $|\eta| > |\delta|$, and (3.6) implies that $|\lambda| < 1$. This concludes the proof of the theorem.

4. ONE MORE SPECIAL CASE

In this section we use the variational characterization of periodic schemes obtained in Sect. 2 to obtain another convergence statement. The angle between m^i and m^j is given by

$$\frac{\langle m^i, m^j\rangle}{\sqrt{\langle m^i, m^i\rangle}\sqrt{\langle m^j, m^j\rangle}} = \frac{a_{i,j}}{\sqrt{a_{i,i}a_{j,j}}}.$$

It is easily seen that if

$$\frac{a_{i,i+1}}{\sqrt{a_{i,i}a_{i+1,i+1}}} < \frac{1}{2},$$

then $D + L_2 + L_2^*$ is positive definite. If, in addition, $a_{1,n} = 0$, then Theorem 3.2 applies. In this section we shall show that if

$$\frac{a_{i,i+1}}{\sqrt{a_{i,i}a_{i+1,i+1}}} < \frac{1}{2}, \tag{4.1}$$

then it is sufficient to have

$$\frac{a_{1,n}}{\sqrt{a_{1,1}a_{n,n}}} < \frac{1}{2} \tag{4.2}$$

for convergence to hold.

We formulate this as Theorem 4.1.

THEOREM 4.1. *If the inequalities* (4.1) *and* (4.2) *are satisfied then the periodic scheme converges.*

Proof. We normalize m^i so that $\|m^i\| = 1$.

We rewrite (2.8) as

$$y^{j+1} = y^j - \alpha^j m^{k_n+1^{(j)}}, \tag{4.1}$$

where

$$\alpha^j = \langle m^{k_n+1^{(j)}}, y^{j-1}\rangle. \tag{4.2}$$

Combining (4.1) and (4.2) gives

$$y^{j+1} = y^j - \langle m^{k_n+1^{(j)}}, y^j + \alpha^{j-1} m^{k_n+1^{(j-1)}}\rangle m^{k_n+1^{(j)}}.$$

If we let $z^j = \alpha^{j-1}$, this relation may be written as

$$z^{j+1} = \langle m^{k_n+1^{(j)}}, y^j\rangle + z^j\langle m^{k_n+1^{(j)}}, m^{k_n+1^{(j-1)}}\rangle. \tag{4.3}$$

Using this expression for α^j $(= z^{j+1})$, (4.1) may be written as

$$\begin{aligned} y^{j+1} = y^j &- (m^{k_n+1^{(j)}}\rangle\langle m^{k_n+1^{(j)}})y^j \\ &- z^j\langle m^{k_n+1^{(j)}}, m^{k_n+1^{(j-1)}}\rangle m^{k_n+1^{(j)}}. \end{aligned} \tag{4.4}$$

Let $w = m^{k_n+1^{(j)}}$ and let $\beta = \langle m^{k_n+1^{(j)}}, m^{k_n+1^{(j-1)}}\rangle = a_{i,i+1}/\sqrt{a_{i,i}a_{i+1,i+1}}$. Then the Eqs. (4.3) and (4.4) may be written in matrix form:

$$\begin{pmatrix} y^{j+1} \\ z^{j+1} \end{pmatrix} = \begin{pmatrix} I - w\rangle\langle w & -\beta w \\ w^T & \beta \end{pmatrix} \begin{pmatrix} y^j \\ z^j \end{pmatrix} \equiv N \begin{pmatrix} y^j \\ z^j \end{pmatrix}. \tag{4.5}$$

We shall now define a quadratic vector norm which, when $|\beta| \leqslant \frac{1}{2}$, makes the map (4.5) contracting. This norm will be seen to be independent of w.

Let $P = I - w\rangle\langle w$, so that $P^2 = P$ and $Pw = 0$. Then N may be written as

$$N = \begin{pmatrix} P & -\beta w \\ w^T & \beta \end{pmatrix}.$$

The quadratic norm which we are seeking is determined by the metric tensor $\begin{pmatrix} aI & 0 \\ 0 & 1 \end{pmatrix}$ with an appropriate scalar a. We first determine a so that

$$N^T \begin{pmatrix} aI & 0 \\ 0 & 1 \end{pmatrix} N - \begin{pmatrix} aI & 0 \\ 0 & 1 \end{pmatrix}$$

is negative semi-definite. We will accomplish this under the hypothesis $|\beta| \leqslant \frac{1}{2}$. Using the properties of P, we have

$$\begin{pmatrix} aI & 0 \\ 0 & 1 \end{pmatrix} - N^T \begin{pmatrix} aI & 0 \\ 0 & 1 \end{pmatrix} N = \begin{pmatrix} (a-1)w\rangle\langle w & -\beta w \\ -\beta w^T & 1 - a\beta^2 - \beta^2 \end{pmatrix}.$$

The quadratic form corresponding to this last matrix is

$$(a-1)\langle w, y\rangle^2 - 2\beta\langle w, y\rangle z + (1 - (a+1)\beta^2)z^2. \tag{4.6}$$

This is nonnegative for all (y, z) iff

$$(a-1), (1 - (a+1)\beta^2), -(4\beta^2 - 4(a-1)(1 - (a+1)\beta^2)) \tag{4.7}$$

are all positive. The last expression in (4.7) is the discriminant of (4.6).

The positivity of this discriminant may be written as:

$$\beta^2 - ((a-1) - (a^2-1))\beta^2) \leqslant 0.$$

This is equivalent to

$$\beta^2 \leqslant \frac{a-1}{a^2}. \tag{4.8}$$

The maximum of $(a - 1)/a^2$ over all a is $\frac{1}{4}$ and is attained at $a = 2$. Thus, $\beta \leqslant \frac{1}{2}$ and $a = 2$ satisfies (4.8). Moreover, these values of a and β also cause the first two expressions in (4.7) to be positive. The proof is completed by a slightly finer argument using the nonsingularity of A, which shows that the product of all the $N - s$ is also (strictly) contracting.

5. A GENERAL THEOREM

As can be seen from the above discussion, not much can be said about convergence in general even for periodic schemes. In order to make statements about the convergence of the full class of chaotic procedures described in Sect. 1, the matrix B must be required to satisfy stronger conditions. In this section, we shall state such conditions. Furthermore, we shall show that these conditions are also necessary, i.e., if they are not satisfied there exists a sequence $\mathscr{S}$ so that the chaotic relaxation scheme $(B, C, \mathscr{S})$ diverges.

In view of the examples, it is surprising that the assumptions which will be required are so weak. For example, if B is defined in terms of A as in Sect. 1, $B = D - E$, then a sufficient condition for convergence is that the Jacobi scheme for the matrix A^+ converge. A^+ is the matrix with entries $a_{i,j}^+$— with

$$a_{i,j}^+ = \begin{cases} a_{i,i}, & i = j \\ -|a_{i,j}|, & i \neq j. \end{cases}$$

Furthermore, if the norm of the Jacobi matrix $I - D^{-1}A^+$ corresponding to A^+ is smaller than 1 a certain amount of overrelaxation depending on the value of the norm is allowed. We shall now state our main result. In what follows the absolute value of a matrix M (or a vector v) will be the matrix $|M|$ (or vector $|v|$) whose components are the absolute value of M (or v). A vector v will be said to be positive if each of its components is positive.

THEOREM. *Let $(B, C, \mathscr{S})$ be a chaotic relaxation scheme.*

(*a*) *The sequence of iterates x_j converges if there exists a positive vector v and a number α, $\alpha < 1$ such that*

$$|B|v \leqslant \alpha v$$

(*b*) *This happens if the spectral radius of $|B|$, $\rho(|B|) < 1$.*

(*c*) *If no such v exists then there exists a sequence $\mathscr{S}_0$ depending on B for which the iterates corresponding to $(B, C, \mathscr{S}_0)$ do not converge.*

Proof. (a) The proof of the sufficiency condition is divided into three parts:

(i) The set $\{\omega | |\omega| \leqslant v\}$ (a rectangular parallelopiped) is contracted into itself by a factor α under multiplication by B.

(ii) The sequence of iterates x^j is bounded.

(iii) The sequence of iterates tends to zero.

(i) Suppose $|\omega| \leqslant v$. Then

$$|B\omega| \leqslant |B|\,|\omega| \leqslant |B|v \leqslant \alpha v.$$

(ii) We shall now show that if

$$|x^{j_1+1}|, \ldots, |x^{j_1+s}| \leqslant av$$

for some constant $a > 0$ and a given j_1 then $|x^j| \leqslant av$ for all $j \geqslant j_1 + s$. (Here s is the parameter which occurs in the definition of a chaotic scheme.) This may be seen by induction on j. For suppose $|x^j| \leqslant av$ for all $j \leqslant j_0$, $j_0 > j_1 + s$. Then either $p \neq k_{n+1}(j_0)$, in which case $|x_p^{j_0+1}| \equiv |x_p^{j_0}| \leqslant av$, or $p = k_{n+1}(j_0)$, in which case

$$x_p^{j_0+1} = \langle b^p, z \rangle,$$

where $z_l = x_l^{j_0 - k_l(j_0)}$ and therefore $|z_l| \leqslant av_l$ for each l, $1 \leqslant l \leqslant n$. But then by the contracting property of B (with v replaced by av), we have

$$\langle b^p, z \rangle \leqslant \alpha a v_p$$

$\therefore$ $x_p^{j_0+1} \leqslant av_p$ for all p. This completes the proof of part (ii).

(iii) To conclude the proof of the sufficiency condition (a), we may note that if $|x^{j_1+1}|, \ldots, |x^{j_1+s}| \leqslant av$ then by what was said above $|x^j| \leqslant av$ for all $j > j_1$. If x_p is updated at time j_2, $j_2 > j_1$, we have for all $j \geqslant j_2$

$$|x_p^j| = \langle b^p, z \rangle$$

for some z with components $z_i \leqslant av_i$. Hence,

$$|x_p^j| \leqslant \alpha a v_p.$$

Let j_3 be an instant so that for each $1 \leqslant i \leqslant n$ there exists j, $j_0 + s \leqslant j \leqslant j_3$ with $k_{n+1}(j) = i$. From what was said above, $|x^j| \leqslant \alpha a v$ for all $j \geqslant j_3$. In particular $|x^j| \leqslant \alpha a v$ for $j_3 + 1 \leqslant j \leqslant j_3 + s$. Repeating the argument above, and using the fact that each component is updated infinitely often, we may observe that if $|x^j| < \alpha^i v$ for all $j \geqslant j_i$ for some j_i there exists j_{i+1} with $|x^{j_{i+1}}| < \alpha^{i+1} v$ for all $j \geqslant j_{i+1}$. This implies that $|x^j| \to 0$ and completes the proof of sufficiency.

(b) We shall first show that if a matrix F has the form

$$F = \begin{pmatrix} F_{11} & F_{12} \\ 0 & F_{22} \end{pmatrix}$$

where for some positive v_1 and v_2 $F_{11}v_1 \leqslant \alpha_1 v_1$, $F_{22}v_2 \leqslant \alpha_2 v_2$. Then for any $\varepsilon > 0$ there exists v with $Fv \leqslant (\alpha + \varepsilon)v$, $\alpha = \max(\alpha_1, \alpha_2)$. Indeed, let $v = (v_1, \gamma v_2)$. Then

$$Fv \leqslant (\alpha_1 v_1 + \gamma F_{12} v_2, \gamma \alpha_2 v_2).$$

By choosing γ sufficiently small it is easy to see how $\alpha_1 v_1 + \gamma F_{12} v_2$ can be made smaller then $(\alpha_1 + \varepsilon)v_1$, while it is always true that $\gamma\alpha_2 v_2 \leqslant \gamma \alpha v_2$. To complete the proof of (b) we wish to show that if $\rho(|B|) < 1$ there exists $v > 0$ with $|B|v < \alpha v$, $\alpha < 1$. This follows by induction on the number of components in the normal form of $|B|$, and the auxilliary result just established above. Indeed, this is true if $|B|$ is irreducible ([4], p. 30). Suppose it is shown for any $|B|$ whose normal form has n components

$$|B| = \begin{pmatrix} B_{11} & B_{12} & B_{1n} \\ & \ddots & \vdots \\ 0 & & B_{1n} \end{pmatrix}.$$

Now suppose that $|B|$ has $n + 1$ components. We write $|B|$ in the form

$$|B| = \begin{pmatrix} F_{11} & F_{12} \\ 0 & F_{22} \end{pmatrix}$$

where F_{22} has n components and F_{11} is irreducible. From this it follows that $F_{11}v_1 \leqslant \alpha_1 v_1$. Furthermore, by the induction hypothesis $F_{22}v_2 \leqslant \alpha_2 v_2$ when $\alpha_1 = \rho(|B|) < 1$ and $\alpha_2 \leqslant 1$. Hence, for some $\gamma > 0$, $v = (v_1, \gamma v_2)$,

$$|B|v \leqslant \alpha v, \qquad \alpha < 1.$$

Conversely, if $|B|v \leqslant \alpha v$, $v > 0$, $\alpha < 1$, B is a contraction in the norm defined by $\|x\| = \max(|x_i|/v_i)$. Hence, the spectral radius of $|B|$ is smaller than 1. This completes the proof of (b).

(c) Suppose there exists no v with $|B|v < v$. Then by (b) $p(|B|) \geqslant 1$.

Let v be the eigenvector of $|B|$ corresponding to the largest real eigenvalue. Then v may be chosen to be real ([4], p. 76) and nonnegative. Using this fact we shall now show how to construct a sequence $\mathscr{S}_0$ so that the sequence of iterates corresponding to the scheme $(B, C, \mathscr{S}_0)$ diverges for some initial state z.

Let z be such that

$$Bz = z + v \qquad \text{or} \qquad (B - I)z = v.$$

If B, C is consistent with an invertible matrix A, then z exists. $\mathscr{S}_0$ is now constructed as follows:

For $1 \leqslant j \leqslant n$ we let

$$k_i(j) = j - 2, \qquad i = 1, \ldots, n, \qquad k_{n+1}(j) = j - 1.$$

If the iteration is started at an initial state z, the following sequence of iterates is produced:

$$z \quad \begin{matrix} z_1 + v_1 \\ z_2 \\ \vdots \\ z_n \end{matrix} \quad \begin{matrix} z_1 + v_1 \\ z_2 + v_2 \\ z_z \\ z_n \end{matrix} \quad \cdots \quad \begin{matrix} z_1 + v_1 \\ z_2 + v_2 \\ \vdots \\ z_n + v_n \end{matrix}$$

where the fact that $\langle b_i, z \rangle = z_i + v_i$ was used repetitively. This sequence may be written as the sum of two sequences $\sigma_1{}^1$, and $\sigma_2{}^1$,

$$z + \frac{v}{2} \quad z_x + \frac{v}{2} \quad , \ldots, \quad z + \frac{v}{2} = \sigma_1{}^1$$

$$\left. \begin{matrix} -\dfrac{v}{2} & \dfrac{v_1}{2} & & \dfrac{v_1}{2} \\ & -\dfrac{v_2}{2} & , \ldots, & \vdots \\ & -\dfrac{v_n}{2} & & \dfrac{v_n}{2} \end{matrix} \right\} = \sigma_2{}^1.$$

Now let $k_{n+1}(n + 1) = 1$, and select $k_i(n + 1)$, $i = 1, \ldots, n$, so that $\operatorname{sgn}(x_i^{n+1-k_i(n+1)}) = \operatorname{sgn} b_i{}^1$. It follows that

$$x_1^{n+2} = \sum_i |b_i{}^1| \, |x_i^{n+1-k_i(n+1)}| = \langle b^1, v \rangle = \rho v_1,$$

where ρ is the spectral radius of $|B|$. We may continue this process, letting $k_{n+1}(j+1) \equiv j \bmod n$ and again choosing $1 \leqslant k_i(j) = n+1$ so that $\operatorname{sgn}(x_i^{j-k_i(j)}) = \operatorname{sgn} b_i^j$ for $n+1 \leqslant j < 2n+1$ and $\operatorname{sgn}(x_i^{j-k_i(j)}) = -\operatorname{sgn} b_i^j$, for $2n+1 \leqslant j < 3n+1$. The result of this iteration procedure up to time $3n+1$ is:

$$(\sigma_1^{\,1}, \sigma_1^{\,2}, \sigma_1^{\,3}) + (\sigma_2^{\,1}, \sigma_2^{\,2}, \sigma_2^{\,3}).$$

Where $\sigma_1^{\,2}$ and $\sigma_1^{\,3}$ are the sequence of iterates generated by a sequence of updatings identical to the one described above but applied to $\sigma_1^{\,1}$, and $\sigma_2^{\,2}$ and $\sigma_2^{\,3}$ have the form:

$$\sigma_2^{\,2} = \begin{array}{cccc}
\dfrac{+\rho v_1}{2} & \dfrac{\rho v_1}{2} & & \dfrac{\rho v_1}{2} \\[2ex]
\dfrac{v_2}{2} & \dfrac{\rho v_2}{2} & & \\[2ex]
 & \dfrac{v_3}{2} & \cdots & \vdots \\[2ex]
\vdots & & & \\[2ex]
 & \vdots & & \\[2ex]
\dfrac{v_n}{2} & \dfrac{v_n}{2} & & \dfrac{\rho v_n}{2}
\end{array}$$

and

$$\sigma_2^{\,3} = \begin{array}{cccc}
\dfrac{-\rho v_1}{2} & \dfrac{-\rho v_1}{2} & & \dfrac{-\rho v_1}{2} \\[2ex]
\dfrac{\rho v_2}{2} & \dfrac{-\rho v_2}{2} & \cdots & \\[2ex]
 & \dfrac{\rho v_3}{2} & & \vdots \\[2ex]
\vdots & & & \\[2ex]
 & \vdots & & \\[2ex]
\dfrac{\rho v_n}{2} & \dfrac{\rho v_n}{2} & & \dfrac{-\rho v_n}{2}
\end{array}.$$

Note now that the last column of $\sigma_2^{\,2}$ together with $\sigma_2^{\,3}$ form an array which is identical with $\sigma_2^{\,1}$ but with $-\rho v_i$ replacing v_i. If we now let

$k_i(j+2n)=k_i(j)$, $i=1,\ldots,n+1$, $j\geqslant n+1$ it is easy to see that the resulting sequence $\sigma_2{}^i$ cannot converge if $\rho\geqslant 1$ and must diverge if $\rho>1$. Suppose now that the sequence $\sigma_1{}^i$ converges. Then $\sigma_1{}^i+\sigma_2{}^i$ is a chaotic relaxation sequence with the initial state z which diverges. If on the other hand $\sigma_1{}^i$ diverges it is easy to see how $\mathscr{S}$ may be modified to produce a sequence

$$z+\frac{v}{2},\sigma_1{}^2,\sigma_1{}^3,\ldots$$

with initial starting state $z+v/2$ which diverges. This concludes the proof of (c) and the theorem.

It may be noted that in proving part (c) of the theorem a great deal of freedom was used in selecting $\mathscr{S}_0$. Clearly weaker conditions could be sufficient to guarantee convergence of a smaller class of chaotic schemes then the full class described by Definition 1. Any finer classification of chaotic schemes yielding successively stronger convergence results would certainly be of some interest.

COROLLARY. *Let the matrix A be written in the form $A=D-E$ when D is a diagonal matrix with elements $d_{i,i}=a_{i,i}$ and $e_{i,j}=a_{i,j}$, $i\neq j$. If $B^\omega=I-\omega D^{-1}A$, $C=\omega D^{-1}$, the scheme $(B^\omega, C^\omega, \mathscr{S})$ converges for all $\mathscr{S}$ satisfying on assumptions of Definition 1 and $\omega=1$ if $\rho(|B|)=\alpha<1$. If $0<\omega<2/(1+\alpha)$ the scheme must converge as well.*

Proof. We must show that $\rho(|B^\omega|)<1$ or alternatively that there exists $v>0$ so that $|B^\omega|v\leqslant\beta v$, $\beta<1$. By (b) of the theorem we know there exists $v>0$ so that $|B^1|v\leqslant\alpha v$. But then

$$|B^\omega|v\leqslant(I(1-\omega)+\omega|B^1|)v\leqslant|(1-\omega)|v+\omega\alpha v=(|1-\omega|+\omega\alpha)v.$$

Let $\beta=(|1-\omega|+\omega\alpha)$. It remains to show that $\beta<1$. If $1\leqslant\omega\leqslant 2/(1+2)$, $\beta=\alpha\omega+(\omega-1)=(1+\alpha)\omega-1<1$. Also if $0\leqslant\omega\leqslant 1$, $\beta=\alpha\omega+(1-\omega)=(-(1-\alpha)\omega+1)<1$ since $\alpha<1$. This completes the proof.

In conclusion, we may note that it follows from the above theorem that if A satisfies one of the following conditions, convergence still holds.

(1) A symmetric and strictly diagonally dominant ([4], p. 23).

(2) A irreducibly diagonally dominant ([4], p. 23).

(3) A symmetric positive definite with nonpositive offdiagonal entries.

We may also note that if in case (3) the spectral radius of $D^{-1}L$ is greater than one, the iteration process is unstable. This can be shown by simply turning the proof of the theorem around. In that case our theorem is as strong as we may wish it to be. If, however, the elements of L are not all negative, it is not clear what happens.

REFERENCES

1 W. Kahan, private communication.
2 P. Stein, Some general theorems on iterants, *J. Res. Natl. Bur. Standards* **48**(1952).
3 J. Rosenfeld, A case study on programming for parallel processors, Research Report RC-1864, IBM Watson Research Center, Yorktown Heights, New York, 1967.
4 R. Varga, *Matrix Iterative Analysis*, Prentice-Hall, Englewood Cliffs, New Jersey, 1962.
5 S. Schechter, Relaxation methods for linear equations, *Comm. Pure and Appl. Math.* **12**(1959), 313–335.
6 A. Ostrowski, Determinanten mit überwiegender Hauptdiagonale und die absolute Konvergenz von linearen Iterationsprozessen, *Comm. Math. Helv.* **30**(1955), 175–210.

Received June 4, 1968